TUGAS MINGGU KETIGA

STATISTIKA DESKRIPTIF

NAMA   : **MUKHAMAD IKHSANUDIN**

NIM   : **082011633086**

S1 SISTEM INFORMASI

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS AIRLANGGA

2021

1. generate data sebanyak 100 ribu

   ```
   MyData <- sample(1:100, 100000, replace =TRUE)
   ```

   a. kuantil, desil, persentil

      - Perintah yang sama dengan argument yang berbeda

        ```
        Kuartil_A <- quantile(NamaData)
        ```

        ```
        > Kuartil_A <- quantile(MyData)
        > Kuartil_A
          0%  25%  50%  75% 100%
           1   25   50   75  100
        ```

        ```
        Desil_A <- quantile(NamaData, c(0:10)/10)
        ```

        ```
        > Desil_A <- quantile(MyData, c(0:10)/10)
        > Desil_A
          0%  10%  20%  30%  40%  50%  60%  70%  80%  90% 100%
           1   10   20   30   40   50   60   70   80   90  100
        ```

        ```
        Persentil_A <- (NamaData, c(0;100)/100)
        ```

        ```
        > Persentil_A <- quantile(MyData, c(0:100)/100)
        > Persentil_A
          0%   1%   2%   3%   4%   5%   6%   7%   8%   9%  10%  11%  12%  13%  14%  15%  16%  17%  18%  19%  20%
           1    2    3    3    4    5    6    7    8    9   10   11   12   13   14   15   16   17   18   19   20
         21%  22%  23%  24%  25%  26%  27%  28%  29%  30%  31%  32%  33%  34%  35%  36%  37%  38%  39%  40%  41%
          21   22   23   24   25   26   27   28   29   30   31   32   33   34   35   36   37   38   39   40   41
         42%  43%  44%  45%  46%  47%  48%  49%  50%  51%  52%  53%  54%  55%  56%  57%  58%  59%  60%  61%  62%
          42   43   44   45   46   47   48   49   50   51   52   53   54   55   56   57   58   59   60   61   62
         63%  64%  65%  66%  67%  68%  69%  70%  71%  72%  73%  74%  75%  76%  77%  78%  79%  80%  81%  82%  83%
          63   64   65   66   67   69   69   70   71   72   73   74   75   76   77   78   79   80   81   82   83
         84%  85%  86%  87%  88%  89%  90%  91%  92%  93%  94%  95%  96%  97%  98%  99% 100%
          84   85   86   87   88   89   90   91   92   93   94   95   97   97   98  100  100
        ```

      - Perintah beda

        ```
        kuartil <- function(x){
         kuartil.x <- quantile(x, seq(0, 1,0.25))
         kuartil.x
        }
        ```

        ```
        > kuartil <- function(x){
        +  kuartil.x <- quantile(x, seq(0, 1,0.25))
        +  kuartil.x
        + }
        > kuartil(MyData)
          0%  25%  50%  75% 100%
           1   25   50   75  100
        ```

        ```
        desil <- function(x){
         desil.x <- quantile(x, seq(0, 1,0.1))
         desil.x
        }
        ```

        ```
        > desil <- function(x){
        +  desil.x <- quantile(x, seq(0, 1,0.1))
        +  desil.x
        + }
        > desil(MyData)
          0%  10%  20%  30%  40%  50%  60%  70%  80%  90% 100%
           1   10   20   30   40   50   60   70   80   90  100
        ```

```
            persentil <- function(x){
              persentil.x <- quantile(x, seq(0, 1,0.01))
              persentil.x
            }
> persentil <- function(x){
+   persentil.x <- quantile(x, seq(0, 1,0.01))
+   persentil.x
+ }
> persentil(MyData)
  0%   1%   2%   3%   4%   5%   6%   7%   8%   9%  10%  11%  12%  13%  14%  15%  16%  17%  18%  19%  20%
   1    2    3    3    4    5    6    7    8    9   10   11   12   13   14   15   16   17   18   19   20
 21%  22%  23%  24%  25%  26%  27%  28%  29%  30%  31%  32%  33%  34%  35%  36%  37%  38%  39%  40%  41%
  21   22   23   24   25   26   27   28   29   30   31   32   33   34   35   36   37   38   39   40   41
 42%  43%  44%  45%  46%  47%  48%  49%  50%  51%  52%  53%  54%  55%  56%  57%  58%  59%  60%  61%  62%
  42   43   44   45   46   47   48   49   50   51   52   53   54   55   56   57   58   59   60   61   62
 63%  64%  65%  66%  67%  68%  69%  70%  71%  72%  73%  74%  75%  76%  77%  78%  79%  80%  81%  82%  83%
  63   64   65   66   67   69   69   70   71   72   73   74   75   76   77   78   79   80   81   82   83
 84%  85%  86%  87%  88%  89%  90%  91%  92%  93%  94%  95%  96%  97%  98%  99% 100%
  84   85   86   87   88   89   90   91   92   93   94   95   97   97   98  100  100
```

b. rentang data dan deviasi standard

```
> max(NamaData) – mix(NamaData)
```

```
> Rentang <- max(MyData)-min(MyData)
> Rentang
[1] 99
```

```
> sd(NamaData)
```

```
> Deviasi <- sd(MyData)
> Deviasi
[1] 28.88169
```
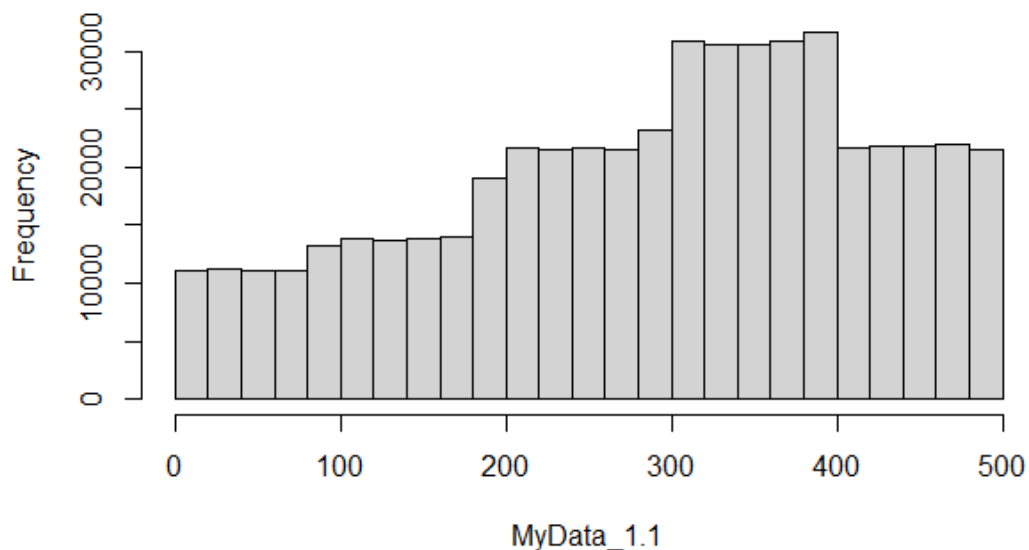
2. generate data sebanyak 500 ribu sedemikian membentuk Kemiringan negative → gunakan histogram

```
> MyData_1.1 <- c(sample(1:90, 50000, replace = TRUE),
sample(90:190, 70000, replace = TRUE), sample(190:300, 120000,
replace = TRUE), sample(300:400, 155000, replace = TRUE),
sample(400:500, 110000, replace = TRUE))
> hist(MyData_1.1)
```
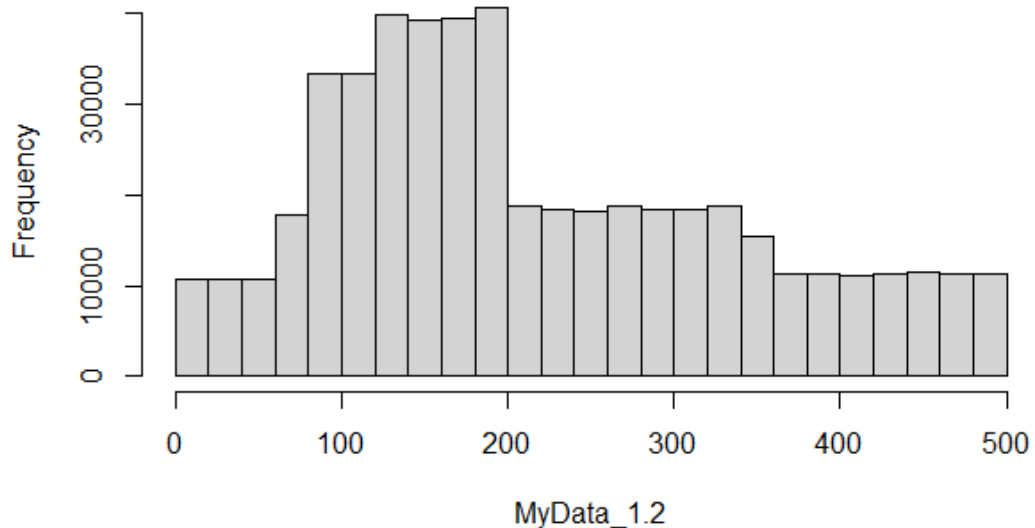

Histogram of MyData_1.1

3. generate data sebanyak 500 ribu sedemikian membentuk Kemiringan positive →
   gunakan histogram

```
> MyData_1.2 <- c(sample(1:75, 40000, replace = TRUE),
sample(75:125, 85000, replace = TRUE), sample(125:200, 150000,
replace = TRUE), sample(200:350, 140000, replace = TRUE),
sample(350:500, 85000, replace = TRUE))
> hist(MyData_1.2)
```
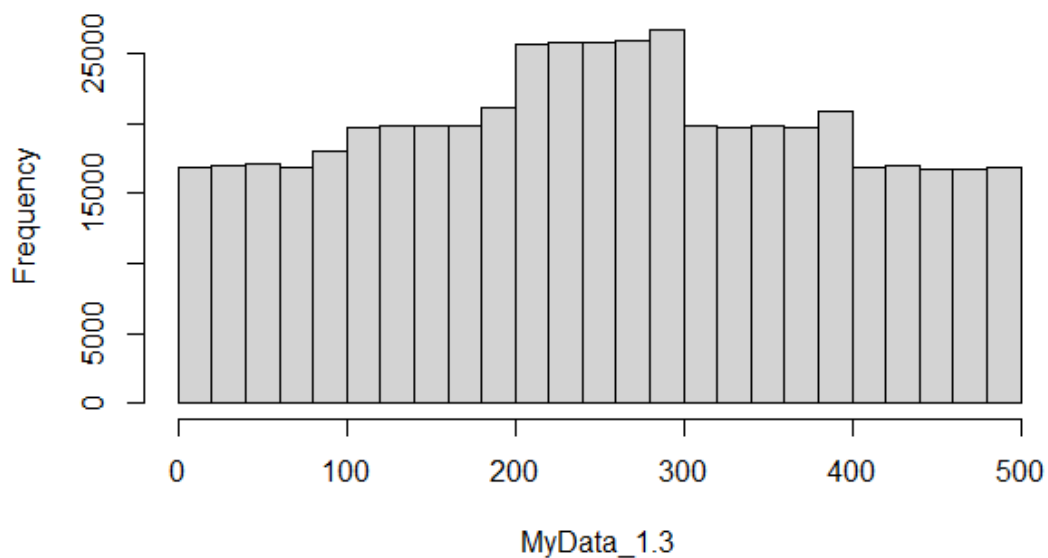
Histogram of MyData_1.2

4. generate data sebanyak 500 ribu Keruncingan negative → gunakan histogram

```
> MyData_1.3 <- c(sample(1:100, 85000, replace = TRUE),
sample(100:200, 100000, replace = TRUE), sample(200:300,
130000, replace = TRUE), sample(300:400, 100000, replace =
TRUE), sample(400:500, 85000, replace = TRUE))
> hist(MyData_1.3)
```

Histogram of MyData_1.3

5. generate data sebanyak 500 ribu Keruncingan positive → gunakan histogram

```
> MyData_1.4 <- c(sample(1:100, 40000, replace = TRUE),
sample(100:210, 60000, replace = TRUE), sample(210:290,
300000, replace = TRUE), sample(290:400, 60000, replace =
TRUE), sample(400:500, 40000, replace = TRUE))
> hist(MyData_1.4)
```

Histogram of MyData_1.4
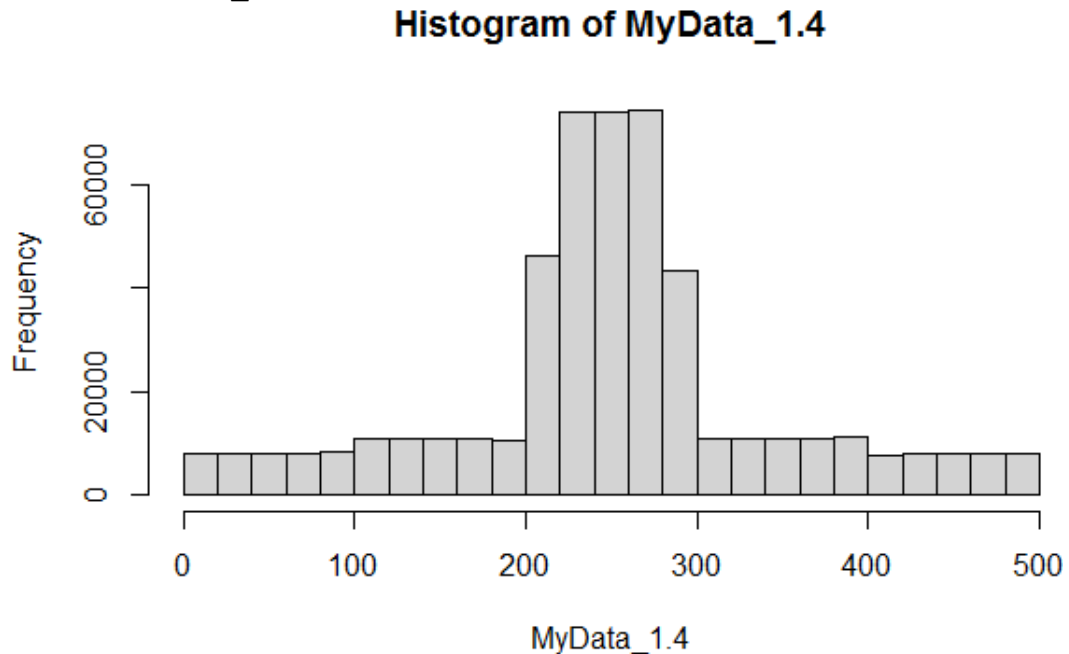
6. untuk no 2 s.d. 5 hitung nilai kemiringan dan keruncingan-nya kemudian jelaskan nilai-nya dan bandingkan dengan gambarnya

- No. 2 (MyData_1.1)

```
> kemiringan = function(data){
  kemiringan = length(data) * (sum(data-mean(data))^3)
/ ((length(data)-1) * (length(data)-2) * (sd(data))^3)
  kemiringan
  }
> kemiringan(MyData_1.1)
[1] -7.234902e-34
```

Kemiringan → Negatif dan sesuai dengan gambar

- No. 3 (MyData_1.2)

```
> kemiringan = function(data){
  kemiringan = length(data) * (sum(data-mean(data))^3)
/ ((length(data)-1) * (length(data)-2) * (sd(data))^3)
  kemiringan
  }
> kemiringan(MyData_1.2)
[1] 2.362199e-33
```

Kemiringan → Positif dan sesuai dengan gambar

- No. 4 (MyData_1.3)

```
> keruncingan = function(data){
+ runcing1 <- (length(data) * (length(data)+1) *
(sum((data-mean(data))^4))) / ((length(data)-1) *
(length(data)-2) * (length(data)-3) * (sd(data))^4)
+ runcing2 <- 3 * ((length(data)-1)^2) /
((length(data)-2) * (length(data)-3))
+ keruncingan = runcing1-runcing2
+ keruncingan
+ }
> keruncingan(MyData_1.3)
[1] -1.020162
```

Keruncingan → Negatif dan sesuai dengan gambar

- No. 5 (MyData_1.4)

```
> keruncingan = function(data){
+ runcing1 <- (length(data) * (length(data)+1) *
(sum((data-mean(data))^4))) / ((length(data)-1) *
(length(data)-2) * (length(data)-3) * (sd(data))^4)
+ runcing2 <- 3 * ((length(data)-1)^2) /
((length(data)-2) * (length(data)-3))
+ keruncingan = runcing1-runcing2
+ keruncingan
+ }
> keruncingan(MyData_1.4)
[1] 0.738092
```
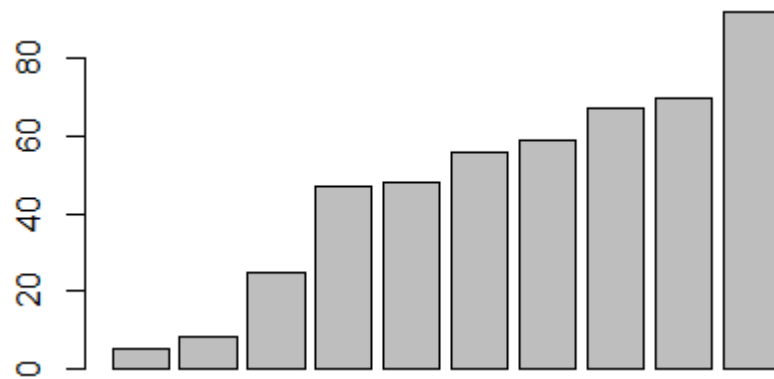
Keruncingan → Positif dan sesuai dengan gambar

7. generate 10 bilangan kemudian buat bar-plot-nya

```
> Bil <- sample (1:100, 10, replace = TRUE)
> Bil
 [1]  5 70 48 47 67 25 56  8 59 92
> barplot(Bil)
```
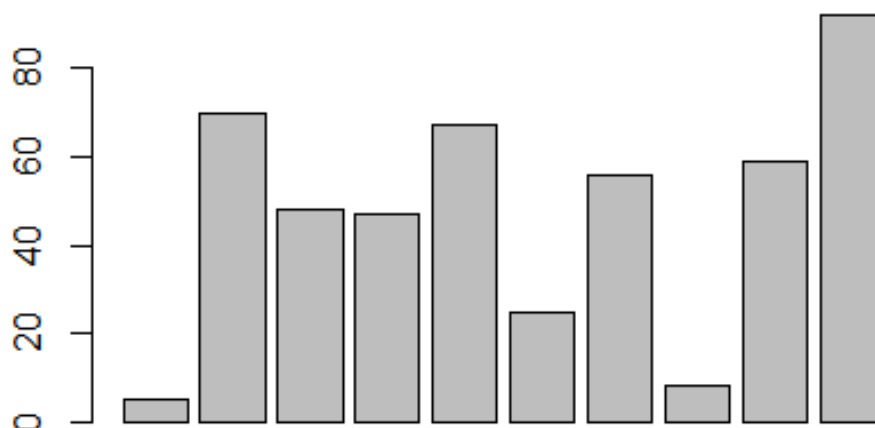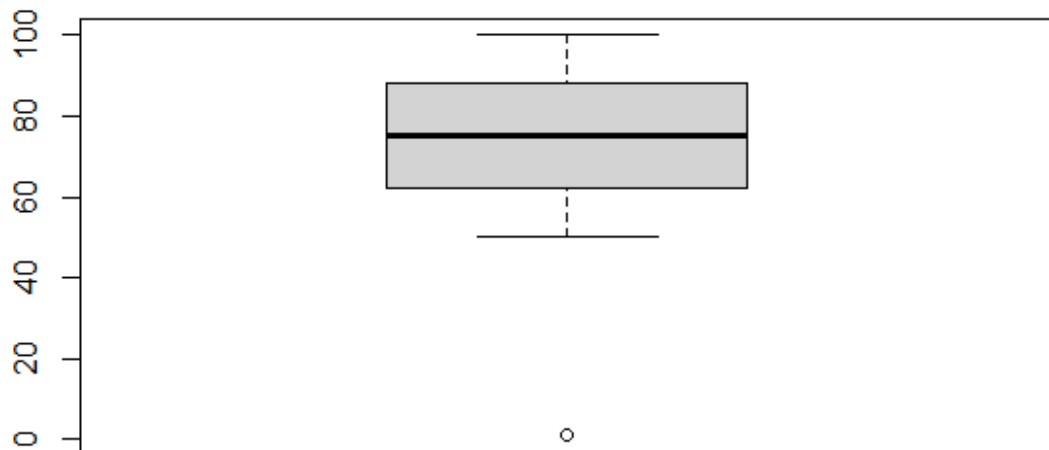
```
> barplot(sort(Bil))
```


8. generate 1000 data yang memuat outlier dan buat box-plot
```
> MyData2 <- c(1, sample(50:100, 999, replace = TRUE))
> boxplot(MyData2
```


9. generate 1000 data yang tidak ada outlier dan buat box-plot
```
> MyData3 <- sample(1:1000, 1000, replace = TRUE)
> boxplot(MyData3)
```